Making Messy Data Work for Conservation

Citation for published version:
Dobson, ADM, Milner-gulland, EJ, Aebischer, NJ, Beale, CM, Brozovic, R, Coals, P, Critchlow, R, Dancer, A, Greve, M, Hinsley, A, Ibbett, H, Johnston, A, Kuiper, T, Le Comber, S, Mahood, SP, Moore, JF, Nilsen, EB, Pocock, MJO, Quinn, A, Travers, H, Wilfred, P, Wright, J & Keane, A 2020, 'Making Messy Data Work for Conservation', *One Earth*, vol. 2, no. 5, pp. 455-465. https://doi.org/10.1016/j.oneear.2020.04.012

Digital Object Identifier (DOI):
10.1016/j.oneear.2020.04.012

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
One Earth

One Earth

Gold Standard for the Global Goals
1 ton of CO_2 offset

CellPress
OPEN ACCESS

Perspective

Making Messy Data Work for Conservation

A.D.M. Dobson,[1,*] E.J. Milner-Gulland,[2] Nicholas J. Aebischer,[3] Colin M. Beale,[4] Robert Brozovic,[5] Peter Coals,[6] Rob Critchlow,[4] Anthony Dancer,[7] Michelle Greve,[8] Amy Hinsley,[6] Harriet Ibbett,[9] Alison Johnston,[10] Timothy Kuiper,[2] Steven Le Comber,[11,20] Simon P. Mahood,[12,13] Jennifer F. Moore,[14] Erlend B. Nilsen,[15] Michael J.O. Pocock,[16] Anthony Quinn,[17] Henry Travers,[2] Paulo Wilfred,[18] Joss Wright,[19] and Aidan Keane[1]
[1]School of GeoSciences, University of Edinburgh, Edinburgh, UK
[2]Department of Zoology, University of Oxford, Oxford, UK
[3]Game & Wildlife Conservation Trust, Fordingbridge, UK
[4]Department of Biology, University of York, York, UK
[5]Frankfurt Zoological Society, Frankfurt, Germany
[6]Wildlife Conservation Research Unit, Department of Zoology, University of Oxford, Oxford, UK
[7]Zoological Society of London, London, UK
[8]Department of Plant and Soil Sciences, University of Pretoria, Pretoria, South Africa
[9]School of Natural Sciences, University of Bangor, Bangor, UK
[10]Department of Ornithology, Cornell University, Ithaca, NY, USA
[11]School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
[12]Wildlife Conservation Society, Cambodia Program, Phnom Penh, Cambodia
[13]Research Institute for the Environment and Livelihoods, Charles Darwin University, Casuarina, NT, Australia
[14]Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL, USA
[15]Department of Terrestrial Ecology, Norwegian Institute for Nature Research, Trondheim, Norway
[16]UK Centre for Ecology & Hydrology, Wallingford, UK
[17]Department of Sociology, Social Policy and Criminology, University of Southampton, Southampton, UK
[18]Department of Life Sciences, The Open University of Tanzania, Dar es Salaam, Tanzania
[19]Oxford Internet Institute, University of Oxford, Oxford, UK
[20]Deceased
*Correspondence: andrew.dobson@ed.ac.uk
https://doi.org/10.1016/j.oneear.2020.04.012

Conservationists increasingly use unstructured observational data, such as citizen science records or ranger patrol observations, to guide decision making. These datasets are often large and relatively cheap to collect, and they have enormous potential. However, the resulting data are generally "messy," and their use can incur considerable costs, some of which are hidden. We present an overview of the opportunities and limitations associated with messy data by explaining how the preferences, skills, and incentives of data collectors affect the quality of the information they contain and the investment required to unlock their potential. Drawing widely from across the sciences, we break down elements of the observation process in order to highlight likely sources of bias and error while emphasizing the importance of cross-disciplinary collaboration. We propose a framework for appraising messy data to guide those engaging with these types of dataset and make them work for conservation and broader sustainability applications.

Challenges and Opportunities

The world's ecosystems face a daunting array of threats, including habitat loss, overexploitation, invasive species, pollution, and climate change.[1–4] Robust data must be the cornerstone for scientists of all stripes seeking to understand the dynamics of environmental change and to map out pathways toward sustainability.[5] Practical decisions for the promotion of environmental health must be evidence based, and conservation interventions are no exception,[6–8] but gathering that evidence via primary data collection within a formal study design is expensive, time consuming, and often impractical.[9,10] Confronted with complex problems and restrictive budgets, governments and conservationists increasingly draw on a large and rapidly growing body of relatively unstructured or semi-structured observational data for monitoring trends and assessing the effect of interventions.[11–13] The use of high-volume, unstructured data has been the subject of a number of recent reviews emphasizing the data-generation potential of social media and other online technologies,[14] the phenomenon of big data,[15,16] and the public understanding of, and participation in, science.[14–18] However, limited attention has been paid to the mechanisms by which problems in such data arise and the ways that these issues may be anticipated (bias avoidance) and overcome (bias mitigation).

Here, we use the umbrella term "messy data" to describe datasets whose collection does not conform to a formal study design and are thus potentially subject to unmeasured bias (Box 1). They are typically generated by processes that are designed either (1) for a separate purpose, wherein the data collection is secondary (e.g., conservation ranger patrols), or (2) for generating the required data but where the observation process is relatively unstructured and/or opportunistic (e.g., many citizen science projects). We use the term "observers" to cover gatherers of any form of messy data, many of whom are unwitting, unpaid, or collecting data as an adjunct to a separate primary objective. Within this definition of "messy" exists a wide range of datasets (Figure 1). For example, in projects such as the Cornell Laboratory of Ornithology's eBird, the survey designers lack

Box 1. A Glossary of Terms

Bias: systematic (as opposed to random) error causing loss of accuracy (as opposed to loss of precision).
Big data: datasets that are too large for traditional data-handling software, as well as typically highly variable data. These data require new methods of storage and analysis to handle the large volumes and tease the signal from the noise.
Citizen science: the intentional, voluntary participation of amateur enthusiasts in scientific research activities. Participants provide data (observational or experimental) and facilities for researchers and may also provide input into project design.
Crowd sensing: the collection of data from large numbers of individuals, each of whom records and submits data on (usually) web-enabled mobile devices such as smartphones.
Distributed mind: describing a complex task split between numerous individuals at the same time, e.g., the protein-folding project, foldit (https://fold.it).
Gamification: the application of game-design elements and game principles in non-game contexts.
Observation process: the many factors that lead to an event being recorded as an observation. This includes the spatial bias of where people are, the chance that the people detect the event, their motivation to record the event, and the accuracy of the record.
Occupancy modeling: an analytical framework designed to explicitly separate the observation process (probability of detection) from the event process (probability of the event), two processes which are otherwise confounded. Typically this modeling framework analyzes binary occurrence data with repeat samples, although a number of extensions allow different data structures.
Semi-structured observational data: data comprising observations made without a standardized observation protocol, as well as important metadata regarding the observation process.
Unstructured observational data: data comprising observations made without a standardized observation protocol.
Web scraping: the extraction of (usually) large amounts of information from online sources, which may or may not occur with the knowledge or permission of the content creator.
Whole-system approach: a method of conducting a project wherein the research question is formulated and investigated with explicit consideration of the full context in which the phenomena of interest, the observations, the analysis, and the responses of interested parties occur.

control over the behavior of observers but have sufficient resources and enough data, metadata, and understanding of the observation process to use sophisticated statistical modeling to account for many aspects of the bias.[20] Other messy datasets, by contrast, contain limited information about the behavior of observers (or other data generators), producing biases that are much harder to tackle. This latter group includes data from herbaria and museums, ranger patrols, illegal wildlife trade seizures at international borders, and crowd-sensing data from social media posts.[21–23]

Messy data have potential advantages over data from structured surveys, including low cost, easy accessibility, high volume, and real-world relevance. In many cases, such data are the only source of information about the phenomenon of interest. For example, assessing past changes in the abundance or distribution of an organism may be impossible without reference to museum records and other historical sources (e.g., Seebens et al.[24] and McClenachan et al.[25]). In other cases, working with data generated for other purposes (e.g., web-scraping listings of wildlife products offered for sale online) allows researchers to study illegal activities without putting themselves in physical danger. However, all messy data have limitations. Any dataset of observations poses three main types of analytical challenge: (1) accounting for errors or mistakes (e.g., incorrect species identification); (2) random variation (or "noise"), which is inherent in the process being observed; and (3) observer bias—systematic errors arising from the observation process (e.g., preferential recording of certain events). In messy data, bias is likely to be especially pervasive, requiring particularly careful consideration.[26–28]

Drawing upon insights from across the natural and social sciences, we synthesize current knowledge and offer guidance for those wishing to engage with messy data. In particular, we challenge the notion that those wishing to use messy data only need to engage with the data after collection. We discuss the importance of weighing, at an early stage, the advantages and disadvantages of using messy data against those of a user-designed, scientifically structured survey. We lay out the steps required to appraise the limitations and potential of a candidate dataset by beginning with an understanding of the underlying observation process—specifically, the way that the data are affected by the motivations, needs, and backgrounds of observers. We illustrate how this exercise serves two purposes: (1) anticipating sources of bias and error and (2) identifying opportunities to align incentives of data users and observers to mutual benefit. Finally, we argue that realizing the full potential of messy data requires researchers and practitioners to adopt a whole-system approach with careful consideration of the entire data life cycle, from problem formulation and data collection to the presentation and use of results.

When Are Messy Data Worth Using?

The global reach of the internet, coupled with the rapid uptake of web-enabled mobile devices, has created unparalleled opportunities to gather low-cost observational data of various types.[29,30] However, although these data may be relatively cheap to acquire, the subsequent cost of collation, appropriate analysis, and interpretation can be high in terms of both time and money; messy data are thus not always worth using.

To take the example of volunteer-collected datasets, organizations managing such projects may require substantial funding to attract, retain, and support volunteers; to maintain data-entry systems; and to validate data.[31] Pocock et al.[31] provide a useful flowchart to guide potential designers of volunteer surveys through the costs and benefits of different types of data,

CellPress
OPEN ACCESS

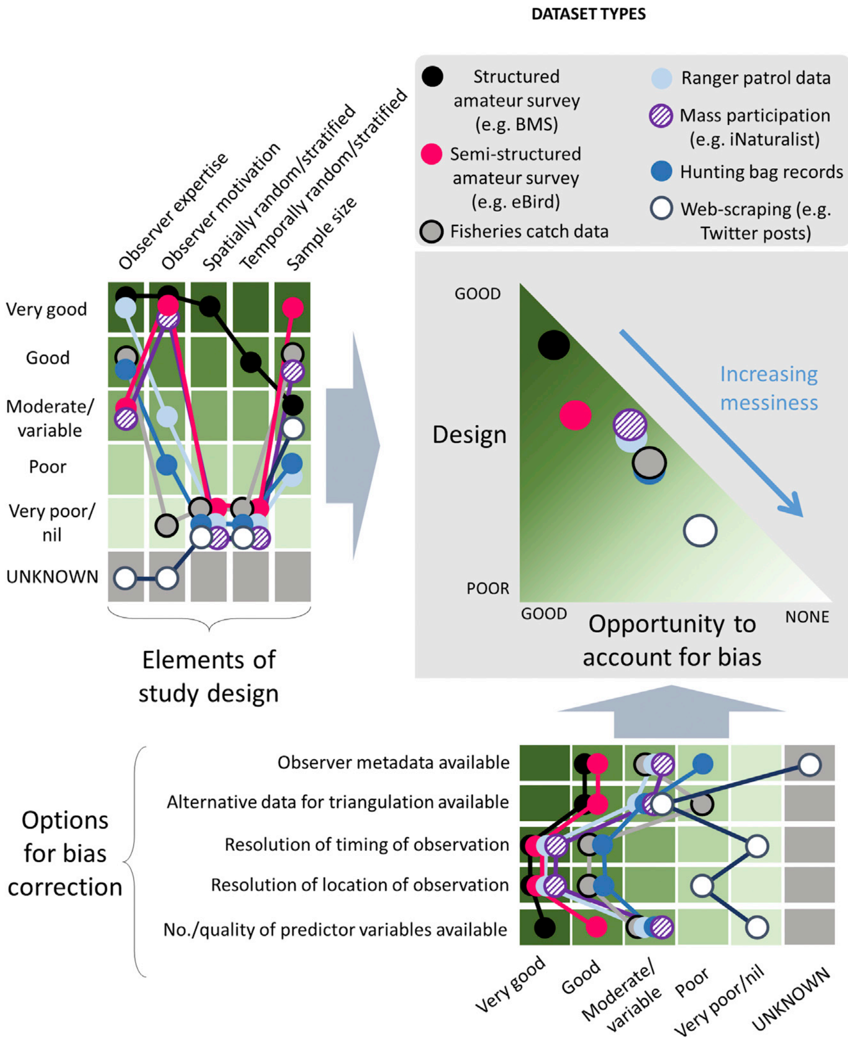

Figure 1. Graphical Illustration of Messy Data

The "messiness" of a dataset can be reduced either prior to its generation via elements of study design (bias avoidance) or afterward via various techniques to improve the information content (bias mitigation). BMS, UK Butterfly Monitoring Scheme.[19]

indicating whether or not the approach is likely to be feasible. For example, UK citizen science projects, such as the Breeding Bird Survey and Butterfly Monitoring Scheme,[19,32] that directly inform policy (e.g., by providing the data for generating biodiversity indicators) cost between £70,000 and £150,000 annually to maintain.[33] They rely upon volunteer observers, but the sampling times, locations, and protocols are nonetheless carefully planned, making the data amenable to analysis and therefore potentially representing good value for money. By contrast, relatively unstructured citizen science data, which generally cost less money to support, may contain less useful information; for example, analysis of DOFbasen, which contains opportunistic sightings of birds in Denmark, showed that it detected fewer than half of the declines in bird population occurrence rates in Denmark apparent in the more structured dataset from the Danish Common Bird Monitoring Scheme.[34] A similar comparison between relatively structured and unstructured datasets for UK birds showed more consistent agreement in trends calculated with simple statistical techniques (90 out of 141 species' trends positively correlated for the two sets), such that agreement was more likely for common and widespread species.[35]

There may also be corollary benefits to set against the potential costs of the messiness of the data produced by volunteer projects, particularly public engagement in science via direct participation.[18] Additionally, the data may have the potential to yield information that might be useful for unanticipated challenges[36] (although this is not exclusive to this type of data). For example, the UK Partridge Count Scheme, a citizen science scheme set up by the Game and Wildlife Conservation Trust to monitor gray partridge (*Perdix perdix*) abundance and breeding success, was subsequently used to evaluate the conservation value of different agri-environmental schemes.[37] Many published studies now include raw data as supplementary material, making data sharing easier. It is also increasingly common for authors to provide code and software, such that future users of messy data may be able to make use of existing methods of data cleaning and processing.

To take another example of a type of increasingly widely used messy data, posts on social media platforms such as Twitter can be searched to provide early warnings of biosecurity risks such as agricultural pests;[38] the costs required to conduct such surveillance with professional observers would be vastly higher and might not necessarily lead to better information. Social media can also illuminate clandestine human behaviors such as illegal wildlife trading (IWT) when the pertinent question is whether this trading is happening, as well as general information on its characteristics rather than detailed questions on trends and absolute magnitudes.[22] However, such data are likely to contain the most pervasive forms of biases while offering very little scope for mitigation. In the case of IWT, variable privacy settings allow some trade to be carried out in relatively public forums, such as open Facebook groups, but vendors will also advertise in closed, private groups or sell directly in private messages. There is therefore no way to know what proportion of the trade is being recorded. Furthermore, although vendors may advertise openly, the sales themselves usually take place in private, meaning that the location and identity of the consumer, or even the final price agreed, may not be known. Careful framing of research questions, together with a good understanding of the data limitations, must be employed before engaging with these types of datasets.[22]

Datasets that originate from social media postings may contain considerable error and bias, but not necessarily as a result of the observation process (which is conducted by researchers). Data collected by non-researcher observers for other conservation purposes are increasingly being used to answer research questions; examples include harvest records from hunters used to develop population management strategies and ranger patrol data used to inform protected area management.[39,40] However, data that cost researchers little or nothing to acquire can nonetheless be expensive to use. In these situations, researchers have relatively little control over the structure of the data that are collected, so biases must be countered during the analysis phase. Indeed, many recent advances in statistics have been driven in large part by the requirement to process large, unstructured datasets.[41] Complex analytical techniques have enabled countless advances across the social and natural sciences and can greatly enhance the utility of observational data. However, complex analysis costs money and has other implications. Firstly, specialist techniques require specialist analysts and software and may require substantial computing time (e.g., Bayesian analysis), all of which have associated barriers to ongoing use (e.g., expertise); open-source software such as R should, however, increase the accessibility of complex analysis. Secondly, statistically characterizing the biases in unstructured data requires a clear understanding of the observation process and appropriate covariate data, which may be either expensive or unavailable. In studies where observers are relatively free to choose the times and places of observations, these factors cannot be accounted for by standardization (which could be achieved with a strict, formalized sampling protocol), making the availability of covariate data especially important. In many such cases, this information will need to have been collected by the observers at the same time as the observations and thus cannot be gathered *post hoc* by subsequent data users. Thirdly, the greater the sophistication of the analysis, the harder it may be to summarize to non-specialist audiences, including existing or potential observers whom one may wish to enthuse and encourage via communication of the outputs.[42] Different audiences require different modes of visualization, and communicating uncertainty in a truthful but accessible way is a challenge.[43] Complex analysis is unlikely to be understood at face value by anyone other than a specialist audience,[44] meaning that poor data-visualization choices can lead to responses ranging from apathy to what has been labeled ''cartohypnosis''—the tendency to invest too much confidence in (spatial) data presented in a suitably authoritative manner.[42,45]

The added value of investments in each of the particular stages of production, processing, and analysis of data could depend upon the lifespan of a project. Whereas statistical analysis can be expensive and time consuming, the benefits of code or software are scalable, so the costs should decline in relative terms as the duration of the project increases. Moreover, the duration of an environmental monitoring study may increase its likelihood of influencing policy.[46]

Even very messy data can sometimes be sufficient to answer questions posed at appropriately low temporal or spatial resolutions or where power to detect change does not need to be high.[47] Sometimes, with limited resources, there may be occasions when it is reasonable simply to use appropriate summaries of raw data without substantial processing to account for bias by acknowledging that biases are likely to be present and being cautious about their interpretation.[48,49] For example, Ingram et al.[48] summarize available data on seizures of pangolin (Pholidota: Manidae) products while drawing attention to the differing availability of data from different sources and without drawing conclusions about the underlying processes.

Overall, judgments about the utility of messy data should be made with reference to a specific objective and should consider the full costs of both collection and analysis given that different questions place different requirements on data quality. Users with restricted budgets should be wary of assuming that large volumes of cheap-to-collect, unstructured data will be better than nothing;[50] the signal-to-noise ratio in unstructured data can be low,[51] and not accounting for biases could lead to misleading conclusions.[52] Therefore, the decision about which datasets to use to answer a question should be taken carefully and deliberately (Figure 2). Before the potential value of a given dataset can be judged, it is necessary to appraise the information it is likely to yield and to identify the best way to extract it. For unstructured or semi-structured observational data, the first step in this appraisal is to consider the observation process itself.

Understanding the Processes Producing the Data

Tolstoy observed that ''happy families are all alike; every unhappy family is unhappy in its own way'';[53] similarly, although standard approaches can be applied to the analysis of relatively structured survey data, the messiest datasets are messy in different ways, and no simple ''recipe'' exists for extracting maximum value from them.

The key to ''reading'' a dataset is to consider exactly what was being recorded, where, when, how, and by whom (Figure 3A). Answering these questions allows the analyst to anticipate the likely sources of bias, which can be subject derived (e.g., heterogeneous detection probabilities), observer derived (e.g., preferential recording of certain events), externally derived (e.g., weather, time of day, or changed instructions), or a combination of all three. For example, electronic healthcare records are routinely collected for patients treated at hospitals and represent a vast store of biomedical information.[54] However, the likelihood that a case of any given medical condition is recorded depends on whether the patient reports to hospital (subject-derived bias) and whether the physician correctly diagnoses the problem and enters the record (observer-derived bias), which in turn is dependent upon the regulatory, policy, and financial environments (externally derived bias).[54,55] Variation in these processes over time or space can confound the underlying phenomenon being measured. For example, apparent changes in depression rates among diabetes and coronary heart disease patients reporting to general practices in Leeds, UK, between 2002 and 2012 were more likely to be driven solely by altered incentives for identifying the condition than by underlying changes in the prevalence of these conditions.[56] An equivalent situation appears to apply to global patterns of ivory seizures; the proportion of ivory transactions seized per country is positively correlated with World Bank governance indicators and most strongly with ''rule of law.''[57]

Determining what is motivating the observers' patterns of behavior can illuminate challenges with interpreting data (Figure 3). For example, birdwatching is a popular and

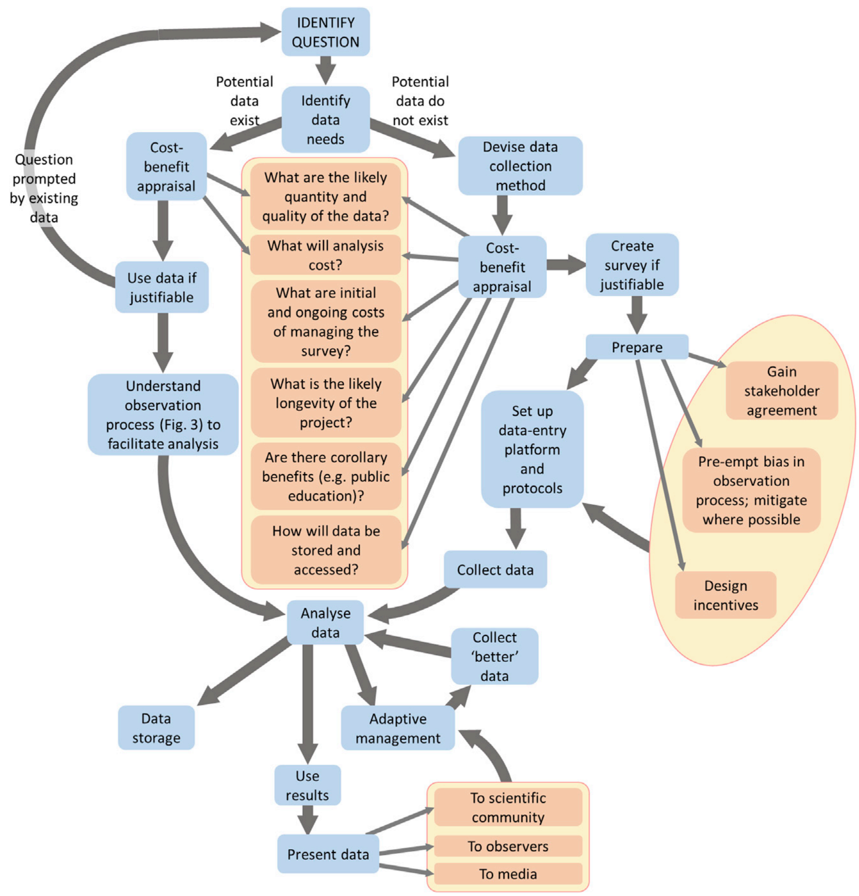

Figure 2. Schematic of Processes Involved in an Adaptive-Management Approach to Using Messy Datasets

widespread pastime, and several organizations have capitalized on this vast potential repository of information; however, projects such as eBird—which aims to use records submitted by members of the public to generate spatial patterns and trends—must contend with the highly non-random probabilities of detection that arise from observers deliberately trying to improve their chances of recording certain species.[58] This taxonomic bias is not restricted to birdwatchers.[59] Even researchers direct their study efforts across taxa in a manner that reflects personal and cultural preferences as much as their relative scientific or conservation importance.[60] In the same way, although ranger patrols may provide data on the abundance and distribution of threats to wildlife, their primary aim is typically to maximize the detection and deterrence of such threats.[39]

The same consideration should be given to the set of external influences that make up the personal, professional, physical, and wider socioeconomic environments of observers.[61] For example, bag data from licensed hunting of large mammals can be influenced by factors that determine the propensity of hunters to fill their quota, including changes in hunting methods and culture, game abundance, and the influence of quotas themselves.[40] Attention should also be paid to the competence of observers in relation to the complexity of observations undertaken;[62] if complex measurements or identification are involved, interobserver variation in technical ability or species recognition skills could produce systematic bias and error.

Once sources of bias in an existing dataset have been identified, there are several options available to counter their influence (i.e., bias mitigation; Figure 3B and Table 1). Statistical modeling is the most familiar of these and is covered in detail elsewhere,[16,27,68] but we further draw readers' attention to the virtual ecologist approach,[69] which involves simulating the underlying ecological data (e.g., species population trends) as well as the observation process used to sample them. The resultant "dummy" data can then be used to test the potential power of different datasets to meet stated objectives and to estimate the degree of analytical complexity that might be required to extract reliable information from a given dataset. For example, Isaac et al.[27] compared models designed to detect temporal biodiversity trends in simulated datasets with known degrees of bias and error, demonstrating that in this case more complex formulations were superior.

Although messy data are typically used only when structured datasets of adequate size cannot be produced, much smaller structured datasets are sometimes available, and these can be used to assess the reliability of outputs from messy datasets via direct comparison.[28,34,70] Alternatively, several features of potentially messy datasets could be used to indirectly infer their quality, such as the existence of iterative design, observer training and testing, and standardization of data input.[62,71] Where applicable, internal consistency—i.e., the degree of agreement between observers—can also be used to reduce error.[72] For example, this form of calibration is used in criminology to contend with systematic differences in reporting rates. Perceptions of police bias can reduce the reporting of crime among certain cohorts[73] and implicit bias in policing can skew arrest rates,[74] leading to under- and over-representation in crime data, respectively. If the extent of these biases can be quantified, biases can be corrected.[75] For instance, so-called "consent searches" by US police officers who suspect an individual of possessing an illegal substance can be used to calculate the "hit rate" for each racial sub-group (i.e., the proportion of searches in which an illegal substance is found). A significantly lower "hit rate" for any given racial group could indicate that targets had been selected on the basis of race as well as of suspicious behavior such that the probability that an individual would be searched differed between groups.[75] Equivalent approaches have been applied to citizen science data, whereby analysts have used interobserver agreement in observations to assign individuals to skill categories, which can be used to calibrate the dataset.[63,76]

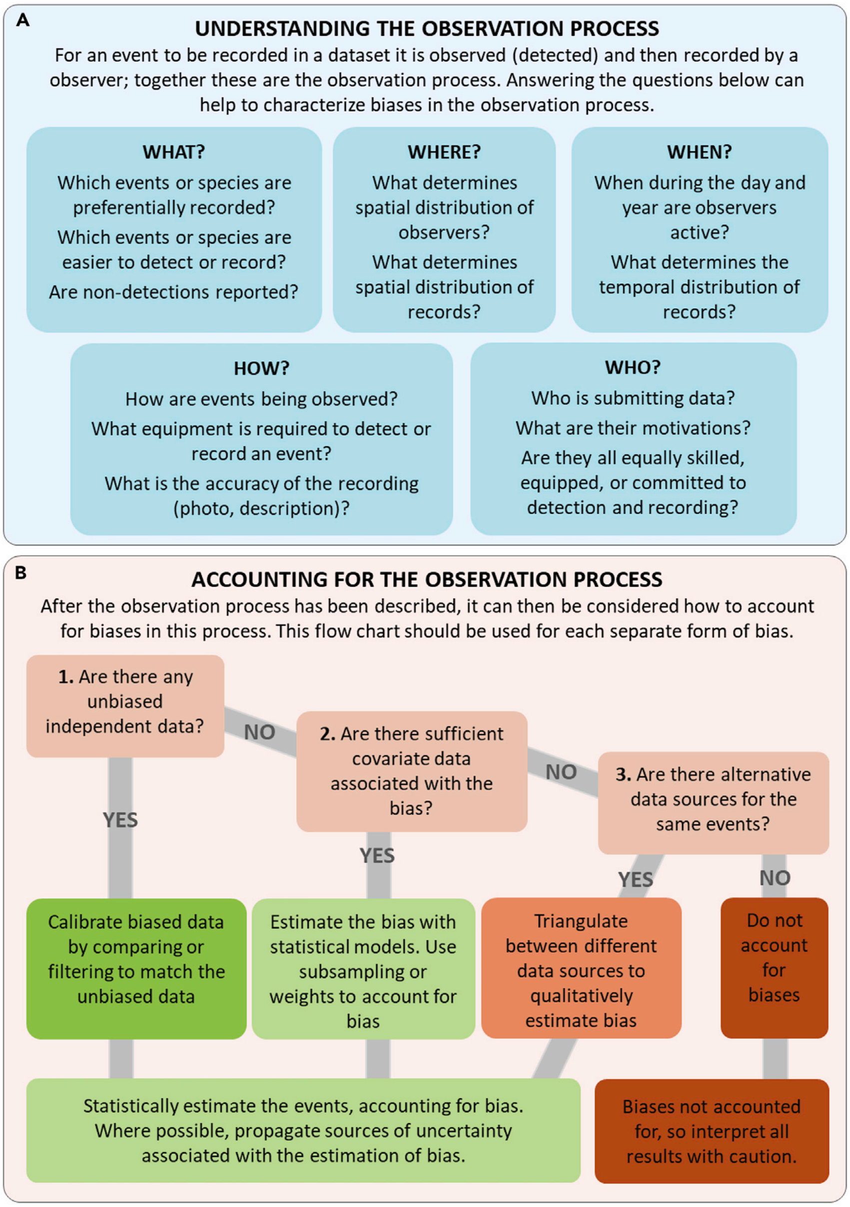

Figure 3. Appraisal and Management of Biases in Observational Data
(A) Questions to ask in order to understand the observation process.
(B) Options for accounting for resultant biases.

Where a messy dataset has not yet been created, study designers may have the opportunity to proactively manipulate the observation process in such a way as to reduce the strength or likelihood of biases before they arise (Figure 2). In the next section, we discuss ways in which this can be achieved.

Aligning Incentives for Planners, Observers, and Users

The collection, analysis, and use of messy datasets typically involve multiple different actors whose motivations may differ. An important but underappreciated avenue for improving the usefulness of messy data for answering practical questions lies in learning what motivates these different actors and designing systems that help to align their incentives. Citizen science projects in particular have utilized a range of techniques to maximize data quality and quantity by aligning the motivations of observers and end users (Table 2). Motivations may be context specific; for example, users of two different "distributed-mind" projects (foldit and GalaxyZoo, designed to find folding solutions for proteins and to classify images of galaxies, respectively) had contrasting views on the desirability of gamification.[91] The needs of end users also differ; for some applications, such as monitoring trends over time, retention of existing observers who carry out repeat visits to the same sites may be more important than the recruitment of new ones.[87] The incentives required to support retention might be quite different from those enhancing recruitment.[92]

In other messy-data contexts, incentive alignment is often less well developed or implemented, and lessons from citizen science may not necessarily be applicable to every context. Where the data gathering is secondary to the main activity undertaken by observers, the potential for aligning incentives may be limited, and for crowd-sensed data there may not be any opportunity at all. Conservation ranger patrols are an example of the former. Here, the primary requirement to find and remove threats to animals will frequently be at odds with the secondary requirement to learn about spatial and temporal patterns in such threats; rangers are very unlikely to perform systematic searches if they already have a perception of where illegal activities are likely to occur. Overall, knowing what, if any,

If several messy datasets are available, it may be possible to mitigate bias via triangulation.[25,77,78] This process is most effective when the datasets originate from different processes and are therefore less likely to share the same sorts of bias.[79,80] For example, Been et al.[81] estimated the use of opioids in Lausanne, Switzerland, by combining public health data with chemical analysis of wastewater. Examples of triangulation with diverse data sources are rare in purely ecological data, although they are relatively common in the conservation social science literature.[82] Therefore, there could be the potential to use this approach more often in conservation contexts, especially if there is active dialog between groups of stakeholders and researchers. For example, the "oakmapper" tool provides a case study of flexibly combining data from public, professional, and regulatory spheres to monitor the spread of a plant pathogen.[83]

Table 1. Conservation Case Studies of Messy Data Use

Example	Challenge	Data	Likely Bias	Solution to Bias	Reference
eBird	generating species distribution estimates	semi-structured, collected by enthusiastic amateurs: eBird, a global volunteer dataset comprising observations of bird species in the form of species checklists	variation in observer experience and behavior leads to unreliable data for less familiar species	average number of species recorded by each observer per checklist is used as a proxy for experience and behavior and used as a covariate in occupancy models	Johnston et al.[63]
Ranger patrols	improving conservation law-enforcement strategy	semi-structured, collected by professionals engaged in alternative activity: ranger-collected observations of illegal activity encountered during the course of patrols, Queen Elizabeth Protected Area, Uganda	observations biased to areas where rangers expect to find illegal activity; lack of spatial evenness	statistical modeling: Bayesian general additive models that explicitly account for imperfect detection (require sufficient sample size in original data)	Critchlow et al.[39,64]
Bird ringing	determining causes of mortality of little owls (*Athene noctua*)	unstructured, collected by any member of the public: capture-mark-recapture of owls ringed in Germany; recaptures are opportune recoveries of marked, deceased individuals by members of the public	ringing recoveries biased to types of mortality (e.g., vehicle collisions) most visible to members of the public	calibrate mortality data by using independent, small-scale telemetry study	Naef-Daenzer et al.[65]
Hunter observations	monitoring population density of moose (*Alces alces*)	poorly structured, collected by enthusiastic amateurs: hunter observations recorded via a smartphone app in Alberta, Canada	lack of spatial information allows duplicate counts; hunters may self-select areas of high moose abundance, thereby providing over-estimates of density	accept bias; calibrate with occasional aerial surveys	Boyce and Corrigan[66]
Wikipedia users	understanding the global cultural attitudes toward reptiles	web scraped: counts of Wikipedia page views of selected reptile species across different language versions	page views likely to be relatively high for species living in areas where internet penetration is high; views restricted by availability of species pages in the language of the potential viewer	accept the existence of biases and interpret with caution	Roll et al.[67]

incentives could be offered requires a good understanding of the motivations of participants.[91] If direct questioning of observers is not possible,[93,94] this might be done by analyzing the composition of recorded data in order to determine which external factors drive particular types of participation.[61,95]

Designers of messy-data-collection programs should also be aware that incentives can be counterproductive. If leader boards or other competition-provoking mechanisms are employed, the "targets" should reflect what the program designer wants. For example, if presence-absence data are required, observers should be rewarded not for what they find but for how much, and where, they have looked.[96] Any sort of stated target, however, risks becoming counterproductive if it encourages the pursuit of a simplified proxy of the actual goal. This broad phenomenon is characterized in public policy as the "cobra effect" after the British Colonial government in India's policy of putting a bounty on dead cobras in an attempt to reduce the populations of these snakes. The populace soon began to breed cobras in order to claim the bounty, thus making the problem worse, not better.[97] As with all interventions, the success of incentives to improve data quantity and quality should be tested and monitored as part of an iterative study design.[62,98,99]

A dataset's size alone is not sufficient to guarantee its utility,[50] and in some instances it may be preferable to incentivize data quality rather than quantity. Callaghan et al.[96] present an approach to increasing the marginal value of ecological observations by estimating the statistical leverage of an observation at a given point in space and time; the method allows program designers to rank times and places where observations would be

Table 2. Examples of Motivations of Observers and Data Users in Citizen Science Projects and Mechanisms Used to Align Them

Project	Activity	Motivation of Observer	Motivation of Data User	Alignment Mechanism	Reference
eBird	spatial mapping of birds	to contribute to science	to achieve even spatial coverage	"avicaching"—assign points to sampling locations, where the value is the inverse of past sampling efforts	Xue et al.[84]
eBird	spatial mapping of birds	to record personal observations	to maximize participation	provide personalized lists, maps, and charts for individual observers	Wood et al.[85]
iNaturalist	spatial mapping of fauna and flora	to compete with other observers	to maximize sample size	publish leader boards on dataset website	Preece[86]
Fishery monitoring in Maine, US	counting migratory fish at pre-determined locations	to record and enter data in a simple manner	to maximize participation	provide simple protocols; use intuitive data-entry systems	Bieluch et al.[87]
The Maine Loon Count	counting breeding pairs of common loons (*Gavia immer*)	to minimize required effort	to retain observers over time	maximize the interval between requested observations	Stockwell and Gallo[88]
Water-quality monitoring projects	collecting biological and chemical data from freshwater and marine sources	to contribute to a personally important objective	to maximize participation	provide feedback to volunteers on the subsequent data uses and even involve them in management decisions	Alender[89]
Old Weather	transcribing weather accounts from historical ship logs	to compete with other observers	to improve accuracy of information	reward loyalty to specific ships, thereby allowing each user to become familiar with the handwriting used on that ship	Eveleigh et al.[90]

most beneficial and to offer this information to prospective observers. This is likely to be most successful when the observers understand the value of their contribution and are motivated to enhance it. Data quality could in some circumstances be improved by better communication of the science behind the study so that observers appreciate why they are being asked to do things in a particular way.[90] For example, coordinators of river herring counts in Maine stressed the need to explain to volunteer counters that zero counts were just as useful and informative as counts where many fish were seen.[87]

More generally, the most successful messy-data projects will be cross-disciplinary collaborations between observers, researchers, analysts, and end users (e.g., policymakers), and it is especially important to maintain effective communication in all directions between each of these components. This form of collaboration serves several purposes, allowing (1) assimilation of local knowledge into wider scientific domains, (2) rapid response to altered circumstances, (3) production of policy-relevant data, and (4) local empowerment for conservation and, overall, therefore, more likelihood that the program will be self-sustaining.[87,88,100] This approach will also promote corollary benefits such as enhanced public understanding of, and engagement with, science. Communication may be most effective when embedded within an adaptive-management approach, wherein feedback loops continuously refine not only the efficiency of the data in finding answers to questions but also the nature of the questions asked (Figure 2). Effective and honest communication not only is useful for achieving a project's aims but should also be considered a minimum ethical standard. Ethics in data collection are beyond the scope of this paper but are nonetheless important and pertinent to many key areas, including privacy, the potentially manipulative effect of incentives, and underlying power dynamics. An ethical code of conduct should be a central component of study design.[101,102]

Concluding Remarks

Messy data have a particularly important role to play in addressing challenges to biodiversity conservation and the broader sustainability agenda, where funding is limited and challenges are seemingly endless. In some cases they represent not merely the best but the only option for gathering information. Where messy datasets originate from citizen science or crowd sensing, sample sizes are typically comparatively large, but the data are likely to contain correspondingly large problems in the form of bias and error (Figure 1). Indeed, the preparation and analysis of messy data frequently requires more money and effort than for more structured datasets, and these relatively hidden costs should be carefully considered before the choice to use messy data is made.

No datasets, however they are generated, are immune to error or bias, and all will require careful analysis. However, the further survey planners depart from "traditional" structured forms of data collection (e.g., moving from small-scale, professional, on-the-ground surveys to methods such as web scraping and global-scale free-to-access data-upload platforms), the greater the problems encountered in analysis. Effort should be invested in identifying techniques that simultaneously maximize data volume while minimizing bias. Gaining a thorough understanding of the observation process, including the motivations and behavior of observers, should be the starting point for anyone wishing to use messy data (Figure 3). Where users have some influence over data collection, taking a broad-scale view of the whole process—from concept to dissemination of results (Figure 2)—

will help to identify the areas where improvements can be most easily and efficiently made.

Increasingly complex analytical techniques will improve the amount of usable information that can be obtained from messy datasets, but this requires skilled staff and sophisticated equipment and could limit transparency and interpretability for non-specialist audiences. Efforts to devise incentives and other protocols that reduce the amount of bias and error entering the data in the first place might be more cost effective. Nonetheless, messy data are both here to stay and hugely valuable for ecological and conservation research if all parties enter into their use in full knowledge of both the benefits and the costs.

ACKNOWLEDGMENTS

This paper is the result of a workshop that took place in Oxford, UK, in March 2019, titled "Learning from observational data to tackle illegal behavior for conservation," and is supported by the Natural Environment Research Council (grant NE/N001370/1).

AUTHOR CONTRIBUTIONS

Conceptualization, all authors; Writing – Original Draft, A.D.M.D., A.K., E.J.M.-G., A.J., and M.G.; Writing – Review and Editing, all authors.

REFERENCES

1. Doherty, T.S., Dickman, C.R., Nimmo, D.G., and Ritchie, E.G. (2015). Multiple threats, or multiplying the threats? Interactions between invasive predators and other ecological disturbances. Biol. Conserv. *190*, 60–68.
2. Young, H.S., McCauley, D.J., Galetti, M., and Dirzo, R. (2016). Patterns, causes, and consequences of anthropocene defaunation. Annu. Rev. Ecol. Evol. Syst. *47*, 333–358.
3. Venter, O., Sanderson, E.W., Magrach, A., Allan, J.R., Beher, J., Jones, K.R., Possingham, H.P., Laurance, W.F., Wood, P., Fekete, B.M., et al. (2016). Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. Nat. Commun. *7*, 12558.
4. Benítez-López, A., Alkemade, R., Schipper, A., Ingram, D., Verweij, P., Eikelboom, J., and Huijbregts, M. (2017). The impact of hunting on tropical mammal and bird populations. Science *356*, 180–183.
5. Cadotte, M.W., Barlow, J., Nuñez, M.A., Pettorelli, N., and Stephens, P.A. (2017). Solving environmental problems in the Anthropocene: the need to bring novel theoretical advances into the applied ecology fold. J. Appl. Ecol. *54*, 1–6.
6. Sutherland, W.J., Pullin, A.S., Dolman, P.M., and Knight, T.M. (2004). The need for evidence-based conservation. Trends Ecol. Evol. *19*, 305–308.
7. Walsh, J.C., Dicks, L.V., and Sutherland, W.J. (2015). The effect of scientific evidence on conservation practitioners' management decisions. Conserv. Biol. *29*, 88–98.
8. Milner-Gulland, E.J., Barlow, J., Cadotte, M.W., Hulme, P.E., Kerby, G., and Whittingham, M.J. (2012). Ensuring applied ecology has impact. J. Appl. Ecol. *49*, 1–5.
9. Gardner, T.A., Barlow, J., Araujo, I.S., Ávila-Pires, T.C., Bonaldo, A.B., Costa, J.E., Esposito, M.C., Ferreira, L.V., Hawes, J., and Hernandez, M.I. (2008). The cost-effectiveness of biodiversity surveys in tropical forests. Ecol. Lett. *11*, 139–150.
10. Vijapure, T., and Sukumaran, S. (2019). Optimization of the taxonomic resolution of an indicator taxon for cost-effective ecological monitoring: perspectives from a heterogeneous tropical coastline. J. Environ. Manage. *247*, 474–483.
11. Follett, R., and Strezov, V. (2015). An analysis of citizen science based research: usage and publication patterns. PLoS One *10*, e0143687.
12. Johnston, A., Fink, D., Reynolds, M.D., Hochachka, W.M., Sullivan, B.L., Bruns, N.E., Hallstein, E., Merrifield, M.S., Matsumoto, S., and Kelling, S. (2015). Abundance models improve spatial and temporal prioritization of conservation resources. Ecol. Appl. *25*, 1749–1756.
13. Woodcock, B.A., Isaac, N.J.B., Bullock, J.M., Roy, D.B., Garthwaite, D.G., Crowe, A., and Pywell, R.F. (2016). Impacts of neonicotinoid use on long-term population changes in wild bees in England. Nat. Commun. *7*, 12459.
14. Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., and Leigh, R.R. (2015). Crowdsourcing for climate and atmospheric sciences: current status and future potential. Int. J. Climatol. *35*, 3185–3203.
15. Mooney, S.J., and Pejaver, V. (2018). Big data in public health: terminology, machine learning, and privacy. Annu. Rev. Public Health *39*, 95–112.
16. Hochachka, W.M., Fink, D., Hutchinson, R.A., Sheldon, D., Wong, W.K., and Kelling, S. (2012). Data-intensive science applied to broad-scale citizen science. Trends Ecol. Evol. *27*, 130–137.
17. Bonney, R., Phillips, T.B., Ballard, H.L., and Enck, J.W. (2016). Can citizen science enhance public understanding of science? Public Underst. Sci. *25*, 2–16.
18. Ballard, H.L., Dixon, C.G.H., and Harris, E.M. (2017). Youth-focused citizen science: examining the role of environmental science learning and agency for conservation. Biol. Conserv. *208*, 65–75.
19. Brereton, T., Botham, M., Middlebrook, I., Randle, Z., Noble, D., and Roy, D. (2017). United Kingdom Butterfly Monitoring Scheme Annual Report 2016 (Centre for Ecology & Hydrology & Butterfly Conservation). https://www.ukbms.org/docs/reports/2016/Butterfly%20Ann%20Report%202016.pdf.
20. Sullivan, B.L., Aycrigg, J.L., Barry, J.H., Bonney, R.E., Bruns, N., Cooper, C.B., Damoulas, T., Dhondt, A.A., Dietterich, T., Farnsworth, A., et al. (2014). The eBird enterprise: an integrated approach to development and application of citizen science. Biol. Conserv. *169*, 31–40.
21. Moore, J.F., Mulindahabi, F., Masozera, M.K., Nichols, J.D., Hines, J.E., Turikunkiko, E., Oli, M.K., and Hayward, M. (2018). Are ranger patrols effective in reducing poaching-related threats within protected areas? J. Appl. Ecol. *55*, 99–107.
22. Hinsley, A., Lee, T.E., Harrison, J.R., and Roberts, D.L. (2016). Estimating the extent and structure of trade in horticultural orchids via social media. Conserv. Biol. *30*, 1038–1047.
23. Runhovde, S.R. (2015). Seizures of inconvenience? Policy, discretion and accidental discoveries in policing the illegal wildlife trade at the Norwegian border. Crime Law Soc. Change *64*, 177–192.
24. Seebens, H., Blackburn, T.M., Dyer, E.E., Genovesi, P., Hulme, P.E., Jeschke, J.M., Pagad, S., Pyšek, P., Winter, M., and Arianoutsou, M. (2017). No saturation in the accumulation of alien species worldwide. Nat. Commun. *8*, 14435.
25. McClenachan, L., Cooper, A.B., McKenzie, M.G., and Drew, J.A. (2015). The importance of surprising results and best practices in historical ecology. BioScience *65*, 932–939.
26. Isaac, N.J., and Pocock, M.J. (2015). Bias and information in biological records. Biol. J. Linn. Soc. *115*, 522–531.
27. Isaac, N.J.B., van Strien, A.J., August, T.A., de Zeeuw, M.P., Roy, D.B., and Anderson, B. (2014). Statistics for citizen science: extracting signals of change from noisy ecological data. Methods Ecol. Evol. *5*, 1052–1060.
28. Aceves-Bueno, E., Adeleye, A.S., Feraud, M., Huang, Y., Tao, M., Yang, Y., and Anderson, S.E. (2017). The accuracy of citizen science data: a quantitative review. Bull. Ecol. Soc. Am. *98*, 278–290.
29. McKinley, D.C., Miller-Rushing, A.J., Ballard, H.L., Bonney, R., Brown, H., Cook-Patton, S.C., Evans, D.M., French, R.A., Parrish, J.K., and Phillips, T.B. (2017). Citizen science can improve conservation science, natural resource management, and environmental protection. Biol. Conserv. *208*, 15–28.
30. Sheil, D., Mugerwa, B., and Fegraus, E.H. (2013). African golden cats, citizen science, and serendipity: tapping the camera trap revolution. South Afr. J. Wildl. Res. *43*, 74–78.
31. Pocock, M.J., Chapman, D.S., Sheppard, L.J., and Roy, H.E. (2014). Choosing and Using Citizen Science: A Guide to When and How to Use Citizen Science to Monitor Biodiversity and the Environment (NERC/Centre for Ecology & Hydrology).
32. Harris, S., Massimino, D., Gillings, S., Eaton, M., Noble, D., Balmer, D., Procter, D., Pearce-Higgins, J., and Woodcock, P. (2018). The Breeding Bird Survey 2017 (British Trust for Ornithology), BTO Research Report 706.
33. Roy, H.E., Pocock, M.J., Preston, C.D., Roy, D.B., Savage, J., Tweddle, J., and Robinson, L. (2012). Understanding Citizen Science and Environmental Monitoring: Final Report on Behalf of UK Environmental Observation Framework (NERC Centre for Ecology & Hydrology/Natural History Museum).
34. Kamp, J., Oppel, S., Heldbjerg, H., Nyegaard, T., and Donald, P.F. (2016). Unstructured citizen science data fail to detect long-term population declines of common birds in Denmark. Divers. Distrib. *22*, 1024–1035.

35. Boersch-Supan, P.H., Trask, A.E., and Baillie, S.R. (2019). Robustness of simple avian population trend models for semi-structured citizen science data is species-dependent. Biol. Conserv. *240*, 108286.

36. Kelling, S., Hochachka, W.M., Fink, D., Riedewald, M., Caruana, R., Ballard, G., and Hooker, G. (2009). Data-intensive science: a new paradigm for biodiversity studies. BioScience *59*, 613–620.

37. Ewald, J., Aebischer, N., Richardson, S., Grice, P., and Cooke, A. (2010). The effect of agri-environment schemes on grey partridges at the farm level in England. Agric. Ecosyst. Environ. *138*, 55–63.

38. Caley, P., Welvaert, M., and Barry, S.C. (2019). Crowd surveillance: estimating citizen science reporting probabilities for insects of biosecurity concern. J. Pest Sci. *93*, 543–550.

39. Critchlow, R., Plumptre, A.J., Alidria, B., Nsubuga, M., Driciru, M., Rwetsiba, A., Wanyama, F., and Beale, C.M. (2017). Improving law-enforcement effectiveness and efficiency in protected areas using ranger-collected monitoring data. Conserv. Lett. *10*, 572–580.

40. Bischof, R., Nilsen, E.B., Brøseth, H., Männil, P., Ozoliņš, J., and Linnell, J.D. (2012). Implementation uncertainty when using recreational hunting to manage carnivores. J. Appl. Ecol. *49*, 824–832.

41. Gandomi, A., and Haider, M. (2015). Beyond the hype: big data concepts, methods, and analytics. Int. J. Inf. Manage. *35*, 137–144.

42. McInerny, G.J., Chen, M., Freeman, R., Gavaghan, D., Meyer, M., Rowland, F., Spiegelhalter, D.J., Stefaner, M., Tessarolo, G., and Hortal, J. (2014). Information visualisation for science and policy: engaging users and avoiding bias. Trends Ecol. Evol. *29*, 148–157.

43. Hamel, P., and Bryant, B.P. (2017). Uncertainty assessment in ecosystem services analyses: seven challenges and practical responses. Ecosyst. Serv. *24*, 1–15.

44. Pidgeon, N., and Fischhoff, B. (2011). The role of social and decision sciences in communicating uncertain climate risks. Nat. Clim. Change *1*, 35.

45. Boggs, S.W. (1947). Cartohypnosis. The Scientific Monthly *64*, 469–476.

46. Hughes, B.B., Beas-Luna, R., Barner, A.K., Brewitt, K., Brumbaugh, D.R., Cerny-Chipman, E.B., Close, S.L., Coblentz, K.E., De Nesnera, K.L., and Drobnitch, S.T. (2017). Long-term studies contribute disproportionately to ecology and policy. BioScience *67*, 271–281.

47. La Sorte, F.A., and Somveille, M. (2019). Survey completeness of a global citizen-science database of bird occurrence. Ecography *43*, 34–43.

48. Ingram, D.J., Cronin, D.T., Challender, D.W.S., Venditti, D.M., and Gonder, M.K. (2019). Characterising trafficking and trade of pangolins in the Gulf of Guinea. Glob. Ecol. Conserv. *17*, e00576.

49. Mendiratta, U., Sheel, V., and Singh, S. (2017). Enforcement seizures reveal large-scale illegal trade in India's tortoises and freshwater turtles. Biol. Conserv. *207*, 100–105.

50. Bayraktarov, E., Ehmke, G., O'Connor, J., Burns, E.L., Nguyen, H.A., McRae, L., Possingham, H.P., and Lindenmayer, D.B. (2019). Do big unstructured biodiversity data mean more knowledge? Front. Ecol. Environ. *6*, 239.

51. Welvaert, M., and Caley, P. (2016). Citizen surveillance for environmental monitoring: combining the efforts of citizen science and crowdsourcing in a quantitative data framework. SpringerPlus *5*, 1890.

52. Buckland, S., and Johnston, A. (2017). Monitoring the biodiversity of regions: key principles and possible pitfalls. Biol. Conserv. *214*, 23–34.

53. Tolstoy, L. (1887). Anna Karenina (Thomas Y. Crowell & Co. Ltd.).

54. Hripcsak, G., and Albers, D.J. (2012). Next-generation phenotyping of electronic health records. J. Am. Med. Inform. Assoc. *20*, 117–121.

55. Foy, R., Leaman, B., McCrorie, C., Petty, D., House, A., Bennett, M., Carder, P., Faulkner, S., Glidewell, L., and West, R. (2016). Prescribed opioids in primary care: cross-sectional and longitudinal analyses of influence of patient and practice characteristics. BMJ Open *6*, e010276.

56. McLintock, K., Russell, A.M., Alderson, S.L., West, R., House, A., Westerman, K., and Foy, R. (2014). The effects of financial incentives for case finding for depression in patients with diabetes and coronary heart disease: interrupted time series analysis. BMJ Open *4*, e005178.

57. Underwood, F.M., Burn, R.W., and Milliken, T. (2013). Dissecting the illegal ivory trade: an analysis of ivory seizures data. PLoS One *8*, e76539.

58. Johnston, A., Hochachka, W., Strimas-Mackey, M., Gutierrez, V.R., Robinson, O., Miller, E., Auer, T., Kelling, S., and Fink, D. (2019). Best practices for making reliable inferences from citizen science data: case study using eBird to estimate species distributions. bioRxiv. https://doi.org/10.1101/574392.

59. Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., and Legendre, F. (2017). Taxonomic bias in biodiversity data and societal preferences. Sci. Rep. *7*, 9132.

60. Jarić, I., Correia, R.A., Roberts, D.L., Gessner, J., Meinard, Y., and Courchamp, F. (2019). On the overlap between scientific and societal taxonomic attentions—insights for conservation. Sci. Total Environ. *648*, 772–778.

61. Boakes, E.H., Gliozzo, G., Seymour, V., Harvey, M., Smith, C., Roy, D.B., and Haklay, M. (2016). Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. Sci. Rep. *6*, 33051.

62. Kosmala, M., Wiggins, A., Swanson, A., and Simmons, B. (2016). Assessing data quality in citizen science. Front. Ecol. Environ. *14*, 551–560.

63. Johnston, A., Fink, D., Hochachka, W.M., Kelling, S., and Isaac, N. (2018). Estimates of observer expertise improve species distributions from citizen science data. Methods Ecol. Evol. *9*, 88–97.

64. Critchlow, R., Plumptre, A.J., Driciru, M., Rwetsiba, A., Stokes, E.J., Tumwesigye, C., Wanyama, F., and Beale, C.M. (2015). Spatiotemporal trends of illegal activities from ranger-collected data in a Ugandan national park. Conserv. Biol. *29*, 1458–1470.

65. Naef-Daenzer, B., Korner-Nievergelt, F., Fiedler, W., and Grüebler, M.U. (2017). Bias in ring-recovery studies: causes of mortality of little owls *Athene noctua* and implications for population assessment. J. Avian Biol. *48*, 266–274.

66. Boyce, M.S., and Corrigan, R. (2017). Moose survey app for population monitoring. Wildl. Soc. Bull. *41*, 125–128.

67. Roll, U., Mittermeier, J.C., Diaz, G.I., Novosolov, M., Feldman, A., Itescu, Y., Meiri, S., and Grenyer, R. (2016). Using Wikipedia page views to explore the cultural importance of global reptiles. Biol. Conserv. *204*, 42–50.

68. Bird, T.J., Bates, A.E., Lefcheck, J.S., Hill, N.A., Thomson, R.J., Edgar, G.J., Stuart-Smith, R.D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J.F., et al. (2014). Statistical solutions for error and bias in global citizen science datasets. Biol. Conserv. *173*, 144–154.

69. Zurell, D., Berger, U., Cabral, J.S., Jeltsch, F., Meynard, C.N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B., et al. (2010). The virtual ecologist approach: simulating data and observers. Oikos *119*, 622–635.

70. Johnston, A., Moran, N., Musgrove, A., Fink, D., and Baillie, S.R. (2020). Estimating species distributions from spatially biased citizen science data. Ecol. Model. *422*, https://doi.org/10.1016/j.ecolmodel.2019.108927.

71. Ratnieks, F.L., Schrell, F., Sheppard, R.C., Brown, E., Bristow, O.E., and Garbuzov, M. (2016). Data reliability in citizen science: learning curve and the effects of training method, volunteer background and experience on identification accuracy of insects visiting ivy flowers. Methods Ecol. Evol. *7*, 1226–1235.

72. Swanson, A., Kosmala, M., Lintott, C., Simpson, R., Smith, A., and Packer, C. (2015). Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. Sci. Data *2*, 150026.

73. Briones-Robinson, R., Powers, R.A., and Socia, K.M. (2016). Sexual orientation bias crimes: examination of reporting, perception of police bias, and differential police response. Crim. Justice Behav. *43*, 1688–1709.

74. Spencer, K.B., Charbonneau, A.K., and Glaser, J. (2016). Implicit bias and policing. Soc. Personal. Psychol. Compass *10*, 50–63.

75. Atiba Goff, P., and Barsamian Kahn, K. (2012). Racial bias in policing: why we know less than we should. Soc. Issues Policy Rev. *6*, 177–210.

76. Horns, J.J., Adler, F.R., and Şekercioğlu, Ç.H. (2018). Using opportunistic citizen science data to estimate avian population trends. Biol. Conserv. *221*, 151–159.

77. Farrell, G., Tseloni, A., and Tilley, N. (2016). Signature dish: triangulation from data signatures to examine the role of security in falling crime. Method. Innov. *9*, https://doi.org/10.1177/2059799115622754.

78. Becken, S., Connolly, R.M., Chen, J., and Stantic, B. (2019). A hybrid is born: integrating collective sensing, citizen science and professional monitoring of the environment. Ecol. Inform. *52*, 35–45.

79. Munafò, M.R., and Smith, G.D. (2018). Robust research needs many lines of evidence. Nature *553*, 399–401.

80. Baker, E., Jeger, M., Mumford, J.D., and Brown, N. (2019). Enhancing plant biosecurity with citizen science monitoring: comparing methodologies using reports of acute oak decline. J. Geogr. Syst. *21*, 111–131.

81. Been, F., Benaglia, L., Lucia, S., Gervasoni, J.-P., Esseiva, P., and Delémont, O. (2015). Data triangulation in the context of opioids monitoring via wastewater analyses. Drug Alcohol Depend. *151*, 203–210.

82. Travers, H., Archer, L.J., Mwedde, G., Roe, D., Baker, J., Plumptre, A., Rwetsiba, A., and Milner-Gulland, E. (2019). Understanding complex

drivers of wildlife crime to design effective conservation interventions. Conserv. Biol. *33*, 1296–1306.

83. Connors, J.P., Lei, S., and Kelly, M. (2012). Citizen science in the age of neogeography: utilizing volunteered geographic information for environmental monitoring. Ann. Assoc. Am. Geogr. *102*, 1267–1289.
84. Xue, Y., Davies, I., Fink, D., Wood, C., and Gomes, C.P. (2016). Avicaching: a two stage game for bias reduction in citizen science. In Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems (International Foundation for Autonomous Agents and Multiagent Systems), pp. 776–785.
85. Wood, C., Sullivan, B., Iliff, M., Fink, D., and Kelling, S. (2011). eBird: engaging birders in science and conservation. PLoS Biol. *9*, e1001220.
86. Preece, J. (2017). How two billion smartphone users can save species! interactions. ACM Interactions *24*, 26–33.
87. Bieluch, K.H., Willis, T., Smith, J., and Wilson, K.A. (2017). The complexities of counting fish: engaging citizen scientists in fish monitoring. Maine Policy Rev. *26*, 9–18.
88. Stockwell, S., and Gallo, S. (2017). Citizen science and wildlife conservation: lessons from 34 Years of the Maine loon count. Maine Policy Rev. *26*, 25–32.
89. Alender, B. (2016). Understanding volunteer motivations to participate in citizen science projects: a deeper look at water quality monitoring. J. Sci. Commun. *15*, A04.
90. Eveleigh, A., Jennett, C., Lynn, S., and Cox, A.L. (2013). "I want to be a captain! I want to be a captain!": gamification in the old weather citizen science project. In Proceedings of the First International Conference on Gameful Design, Research, and Applications, L. Nacke, K. Harrigan, and N. Randall, eds. (ACM), pp. 79–82.
91. Ponti, M., Hillman, T., Kullenberg, C., and Kasperowski, D. (2018). Getting it right or being top rank: games in citizen science. Citiz. Sci. *3*, https://doi.org/10.5334/cstp.101.
92. Selinske, M.J., Coetzee, J., Purnell, K., and Knight, A.T. (2015). Understanding the motivations, satisfaction, and retention of landowners in private land conservation programs. Conserv. Lett. *8*, 282–289.
93. Wright, D.R., Underhill, L.G., Keene, M., and Knight, A.T. (2015). Understanding the motivations and satisfactions of volunteers to improve the effectiveness of citizen science programs. Soc. Nat. Resour. *28*, 1013–1029.
94. Geoghegan, H., Dyke, A., Pateman, R., West, S., and Everett, G. (2016). Understanding Motivations for Citizen Science. Final Report on Behalf of UKEOF (University of Reading, Stockholm Environment Institute, and University of the West of England).
95. Mao, A., Kamar, E., Chen, Y., Horvitz, E., Schwamb, M.E., Lintott, C.J., and Smith, A.M. (2013). Volunteering versus work for pay: incentives and tradeoffs in crowdsourcing. In Proceedings of the First AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2013), 94–102.
96. Callaghan, C.T., Rowley, J.J., Cornwell, W.K., Poore, A.G., and Major, R.E. (2019). Improving big citizen science data: moving beyond haphazard sampling. PLoS Biol. *17*, e3000357.
97. Siebert, H. (2001). Der Kobra-Effekt: Wie man Irrwege der Wirtschaftspolitik vermeidet (Dt. Verlag-Anst).
98. Nichols, J.D., Blohm, R.J., Reynolds, R.E., Trost, R.E., Hines, J.E., and Bladen, J.P. (1991). Band reporting rates for mallards with reward bands of different dollar values. J. Wildl. Manage. *55*, 119–126.
99. Sackett, D.K., and Catalano, M. (2017). Spatial heterogeneity, variable rewards, tag loss, and tagging mortality affect the performance of mark-recapture designs to estimate exploitation: an example using Red Snapper in the northern Gulf of Mexico. N. Am. J. Fish. Manage. *37*, 558–573.
100. Warton, D.I. (2015). New opportunities at the interface between ecology and statistics. Methods Ecol. Evol. *6*, 363–365.
101. Resnik, D.B., Elliott, K.C., Miller, A.K.J.E.S., and Policy. (2015). A framework for addressing ethical issues in citizen science. Environ. Sci. Policy *54*, 475–481.
102. Williamson, V. (2016). On the ethics of crowdsourced research. Polit. Sci. Polit. *49*, 77–81.